



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Establishing propositional truth-value in counterfactual and real-world contexts during sentence comprehension: Differential sensitivity of the left and right inferior frontal gyri

Citation for published version:

Nieuwland, MS 2012, 'Establishing propositional truth-value in counterfactual and real-world contexts during sentence comprehension: Differential sensitivity of the left and right inferior frontal gyri', *NeuroImage*, vol. 59, no. 4, pp. 3433-3440. <https://doi.org/10.1016/j.neuroimage.2011.11.018>

Digital Object Identifier (DOI):

[10.1016/j.neuroimage.2011.11.018](https://doi.org/10.1016/j.neuroimage.2011.11.018)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

NeuroImage

Publisher Rights Statement:

© Nieuwland, M. S. 15 Feb 2012, Establishing propositional truth-value in counterfactual and real-world contexts during sentence comprehension: Differential sensitivity of the left and right inferior frontal gyri, in *NeuroImage*. 59, 4, p. 3433-3440. © Elsevier. The final publication is available at <http://dx.doi.org/10.1016/j.neuroimage.2011.11.018>

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Establishing propositional truth-value in counterfactual and real-world contexts during sentence comprehension: Differential sensitivity of the left and right inferior frontal gyri

Mante S. Nieuwland

Basque Center on Cognition, Brain and Language

Date: 24 October 2011

Keywords: Language comprehension, real-world knowledge, propositional truth-value, counterfactuals, fMRI, LIFG, RIFG

Address correspondence to:

Dr. Mante S. Nieuwland

Basque Center on Cognition, Brain and Language (BCBL)

Paseo Mikeletegi 69-2

20009, Donostia-San Sebastian, Spain

ABSTRACT

What makes a proposition true or false has traditionally played an essential role in philosophical and linguistic theories of meaning. A comprehensive neurobiological theory of language must ultimately be able to explain the combined contributions of real-world truth-value and discourse context to sentence meaning. This fMRI study investigated the neural circuits that are sensitive to the propositional truth-value of sentences about counterfactual worlds, aiming to reveal differential hemispheric sensitivity of the inferior prefrontal gyri to counterfactual truth-value and real-world truth-value. Participants read true or false counterfactual conditional sentences ("If N.A.S.A. had not developed its Apollo Project, the first country to land on the moon would be Russia/America") and real-world sentences ("Because N.A.S.A. developed its Apollo Project, the first country to land on the moon has been America/Russia") that were matched on contextual constraint and truth-value. ROI analyses showed that whereas the left BA 47 showed similar activity increases to counterfactual false sentences and to real-world false sentences (compared to true sentences), the right BA 47 showed a larger increase for counterfactual false sentences. Moreover, whole-brain analyses revealed a distributed neural circuit for dealing with propositional truth-value. These results constitute the first evidence for hemispheric differences in processing counterfactual truth-value and real-world truth-value, and point towards additional right hemisphere involvement in counterfactual comprehension.

INTRODUCTION

Language is a computationally remarkable, uniquely human system, not to mention our principal and most efficient means of communication. It is often assumed that language evolved as an adaptation for communication about the world (e.g., Pinker & Bloom, 1990). The exchange of information through statement of fact routinely affords verification processes through which people are able to agree or disagree with what they read or hear, and draws upon our capacity to recall word meanings and everyday facts and events (i.e., declarative memory; Eichenbaum, 2000). Conditions that make a proposition true or false have traditionally played an essential role in philosophical and linguistic theories of meaning (e.g., Montague, 1973; Tarski, 1944). Yet, while language retains its power for exchanging information about the world, its productive and combinatorial nature enables us to think and talk about concepts beyond the real world. One prominent example of this cognitive ability is counterfactual reasoning about what is in fact false *as if* it were true (e.g., “If China had entered the Vietnam war, then...”), which is pervasive in everyday life (e.g., Byrne, 2002; Kahneman & Miller, 1986; Roese, 1997), and considered one of the hallmarks of complex reasoning skills (e.g., Braine & O’Brien, 1991; Johnson-Laird & Byrne, 2002). Counterfactual comprehension is often thought to require keeping in mind both what is true and what is false (e.g., Byrne, 2002, 2007; for discussion see Evans, 2006; Evans, Over & Handley, 2005). This makes counterfactual comprehension an interesting test-case for identifying the brain regions that are sensitive to propositional truth-value; whereas establishing the truth-value of a regular declarative sentence requires a straightforward mapping of its propositional meaning onto real-world knowledge that is stored in long-term memory, establishing counterfactual truth-value requires the online

construction of a contextually relevant interpretation by temporarily ‘bracketing’ factual knowledge about the real world (e.g., Stalnaker, 1968; Searle, 1975). The current fMRI study aims to unravel the neural circuits that are sensitive to propositional truth-value, as expressed in sentences about counterfactual worlds or about the real world.

According to cognitive theories of text comprehension (e.g., Gernsbacher, 1997; Kintsch, 1988; Myers & O’Brien, 1998), comprehension of counterfactual language requires the suppression or inhibition of automatically activated world knowledge. All information in memory that is related to unfolding linguistic input is initially activated, and this network of co-activations is subsequently pruned so that only the information that is most relevant to the ongoing interpretation is kept active. Similar assumptions have been made in neurocognitive accounts of semantic processing during language comprehension (e.g., Bookheimer, 2002; Hagoort, 2005; Hagoort, Baggio & Willems, 2009; Lau, Phillips & Poeppel, 2008): word-elicited semantic information is selected for message-level semantic integration through the activation of relevant concepts and inhibition of competing concepts, drawing upon activity in and interaction between the temporal and inferior prefrontal brain regions that subserve semantic memory processes (see also Badre & Wagner, 2002; Thompson-Schill, D’Esposito, Aguirre, & Farah, 1997).

A canonical finding supporting the role of the left inferior frontal gyrus for recruiting factual world knowledge during sentence comprehension was reported by Hagoort, Hald, Bastiaansen and Petersson (2004). In an fMRI experiment, sentences that were regarded by Dutch participants as false (e.g., “Dutch trains are white”) elicited common activity increases in the left inferior frontal gyrus (LIFG, BA 45/47) as sentences that contained lexical-semantic anomalies (e.g., “Dutch trains are sour”, which is not only

false but also makes no sense given the semantic constraints of the words ‘trains’ and ‘sour’), each compared to true sentences (e.g., “Dutch trains are yellow”). Later studies replicated and extended these findings by showing that world knowledge violations elicit additional activations in other brain regions, most notably in right inferior prefrontal cortex, middle and superior prefrontal gyrus, middle temporal gyrus, and the inferior parietal lobule (e.g., Groen et al., 2010; Menenti, Petersson, Scheeringa & Hagoort, 2009; Tesink, Buitelaar, Petersson, Van der Gaag, Teunisse & Hagoort, 2011; Tesink, Petersson, Van Berkum, Van den Brink, Buitelaar & Hagoort, 2009). These findings testify to the fact that an extended network of brain regions is sensitive to sentence truth-value, even if participants are not explicitly instructed to establish truth-value.

These findings are consistent with a recent neurocognitive framework of language comprehension, the Memory-Unification-Control framework (e.g., Hagoort, 2005; Hagoort et al., 2009), which posits that the left inferior frontal cortex (with additional support from other areas) has top-down control over the combinatorial semantic processes that compose multi-word utterances from word-elicited information as represented by the left middle temporal cortex. Moreover, knowledge from different sources and modalities (e.g., speaker identity, real-world knowledge, discourse context) is immediately brought to bear on utterance interpretation, by the same brain system that combines the meanings of individual words into a larger whole (Hagoort, 2005; see also Hagoort & Van Berkum, 2007). However, while many event-related potential (ERP) studies have revealed a rapid interaction between world knowledge and discourse context (e.g., Hald, Steenbeek-planting & Hagoort, 2007; Nieuwland & Kuperberg, 2008; Nieuwland & Van Berkum, 2006; Van Berkum, Holleman, Nieuwland, Otten & Murre, 2009; Van Berkum, Van den Brink,

Tesink, Kos & Hagoort, 2008; for eye-tracking and ERP results on online counterfactual comprehension, see Ferguson, Sanford & Leuthold, 2008; Ferguson & Sanford, 2008), the neural mechanisms that underlie the unification of real-world knowledge with discourse context remain understudied. A comprehensive neurobiological theory of language must ultimately be able explain the combined contributions of real-world truth-value and discourse context.

In the only fMRI study thus far that examined these issues (Menenti et al., 2009), sentences that - without any context - were regarded as false or true (e.g., “Donald Duck’s nephews are thieves/boy scouts”) were preceded by either a counterfactual (i.e. fictional) context that was congruent with the false sentences (e.g., a story about early sketches wherein the three nephews were depicted as young bad boys who rob old ladies) or by a neutral context that was congruent with the true sentences (e.g., a story about early sketches where they were depicted as helping old ladies). The results showed that the impact of world knowledge violations (false minus true sentences) was more reduced due to counterfactual context in the RIFG than in the LIFG: world knowledge violations did not elicit effects in RIFG but remained to elicit significant effects in the LIFG despite the counterfactual context. Additionally, in the left angular gyrus, the effect of the world knowledge violation was actually reversed following the counterfactual context, suggesting that this brain region was most sensitive to discourse coherence. The authors offered a tentative interpretation that whereas the LIFG continues to relate incoming information to prior world knowledge, the RIFG is more sensitive to whether incoming information is congruent with the discourse (however, counterfactual context did not reverse the effect of world knowledge in the RIFG). This functional division of labour between LIFG and RIFG

is congruent with a body of literature on right hemisphere involvement in high-level aspects of language (e.g., discourse processing and pragmatic language comprehension; Bookheimer, 2002; Jung-Beeman, 2005; Mason & Just, 2006; Ferstl, Neumann, Bogler & von Cramon, 2008). One influential account holds that the right hemisphere processes incoming information more coarsely and is therefore more sensitive to unusual or novel semantic relationships (e.g., Jung-Beeman, 2005). Semantic selection processes, which inhibit competing concepts to select one concept for message-level semantic integration, may then draw upon the RIFG more strongly when the information that needs to be inhibited is more active in the right hemisphere (e.g., Jung-Beeman, 2005; see also Faust & Gernsbacher, 1996, for discussion). Perhaps in a similar vein, neurocognitive research on reasoning has revealed right lateral prefrontal activations when participants inhibit their world knowledge in order to arrive at a logically sound conclusion (Goel & Dolan, 2003; see also Goel, 2007).

The current fMRI study therefore aimed to address the following question: Are the LIFG and RIFG differentially engaged in balancing the recruitment of information in long-term memory with the online construction of a discourse-relevant and contextualized interpretation? Additionally, and more generally, what brain regions are sensitive to sentence truth-value? This study aims to answer those questions in the context of counterfactual sentence comprehension, by directly comparing the neural processing consequences of sentences that are false with regard to a hypothetical, counterfactual world with those of sentences that are false with regard to the current real-world knowledge. Participants read counterfactual and real-world sentences in which a specific critical word, belonging to a word pair, rendered the sentence false or true. For each word pair, one word

rendered the counterfactual sentence true while rendering the real-world sentence false, and vice versa for the other word (see Table 1 and 2). The contextual constraints were matched for counterfactual and real-world sentences, so that, on average, critical words that rendered sentences true were equally expected for each sentence type, and so that these sentence types were similar in truth-value (as had been established in an independent sentence completion pre-test and truth-value rating pre-test, respectively; see Methods section). The results from an earlier ERP study with these materials (Nieuwland & Martin, 2011) showed that words that rendered sentences false elicited identical N400 effects (compared to words that rendered sentences true) in counterfactual sentences and real-world sentences. These results suggest that, if described consequences are true and predictable given the counterfactual premise, real-world knowledge does not impede or delay counterfactual comprehension, at least at the moment that propositional truth-value can be established. However, fMRI might reveal differences between counterfactual or real-world sentence comprehension processes that occur at a wider timescale and account for the larger network that is engaged for establishing propositional truth-value in both contexts.

Table 1. Example sentences and approximate translations with average truth-value rating and cloze value of the critical word for each condition			
Condition	Example sentences	Mean rating of truth-value	Mean cloze value (%)
Counterfactual True (CT)	<i>Si la N.A.S.A. no hubiera desarrollado su proyecto Apollo, el primer país en pisar la luna habría sido <u>Rusia</u> seguramente.</i> “If N.A.S.A. had not developed its Apollo Project, the first country to land on the moon would have been <u>Russia</u> surely.”	5.62 (.97)	68 (22)
Counterfactual False (CF)	<i>Si la N.A.S.A. no hubiera desarrollado su proyecto Apollo, el primer país en pisar la luna habría sido <u>América</u> seguramente.</i> “If N.A.S.A. had not developed its Apollo Project, the first country to land on the moon would have been <u>America</u> surely.”	1.68 (.74)	-
Real-World True (RWT)	<i>Como la N.A.S.A. desarrolló su proyecto Apollo, el primer país en pisar la luna ha sido <u>América</u> seguramente.</i> “Because N.A.S.A. developed its Apollo Project, the first country to land on the moon was <u>America</u> surely.”	5.50 (1.03)	65 (25)
Real-World False (RWF)	<i>Como la N.A.S.A. desarrolló su proyecto Apollo, el primer país en pisar la luna ha sido <u>Rusia</u> seguramente.</i> “Because N.A.S.A. developed its Apollo Project, the first country to land on the moon was <u>Russia</u> surely.”	1.55 (.64)	-
Filler sentence	<i>La gastronomía vasca está muy bien considerada, es conocida en todo el mundo.</i> “Basque gastronomy is highly regarded, it is known around the world”		
Note. Standard deviations are given in parentheses. Critical words are underlined for expository purposes. For truth-value rating, 1 = False, 7 = True. This table is adapted from Nieuwland & Martin (2011).			

The current study is the first to deal directly with the comprehension of counterfactual conditional sentences. The main prediction of the current study is that the LIFG and RIFG will be differentially sensitive to real-world and counterfactual truth-value: the RIFG being more sensitive to counterfactual truth-value than the LIFG, but being equally or perhaps less sensitive to real-world truth-value (e.g., Hagoort et al., 2004).

An additional, more general objective of this study is to reveal the neural circuits that are sensitive to truth-value, using a design where propositional truth-value of the two sentence types hinges on opposite pairs of critical words, thereby controlling for the impact of differences in lexical-associative factors between words in true and false sentences. Compared to true sentences, false sentences are predicted to increase activation in left and right inferior prefrontal gyrus, and possibly in left middle frontal gyrus, left middle temporal gyrus and in left inferior parietal cortex (e.g., Groen et al., 2010; Hagoort et al., 2004; Menenti et al., 2009; Tesink et al., 2009, 2011).

MATERIALS AND METHODS

Participants

Twenty-eight right-handed students (14 males, mean age = 22.9 years) participated in this study for monetary compensation. All participants were native Spanish speakers, had normal or corrected-to-normal vision, and gave written informed consent. Participants had no neurological or psychiatric disorders, nor had they seen the materials before in an earlier experiment. Four participants were excluded from the final analysis due to excessive movement during the experiment.

Development and Pretest of Materials

Materials were identical to those used by Nieuwland & Martin (2011), and were selected from 133 Spanish sentence quadruplets with two counterfactual and two real-world sentences (see Table 1). Critical words were predicates, nouns or proper names, and never sentence-final. Counterfactual sentences described hypothetical consequences of common-knowledge historical events not having taken place, whereas real-world sentences described actual consequences of these events. The two sentence types differed in three respects: counterfactuals started with the conditional ‘Si’, contained a negative premise, and involved conditional verb tense, real-world sentences started with ‘Como’ (‘because’), were affirmative, and contained no conditional verb tense.

We first established the expectedness of critical words. Twenty students of the University of the Basque Country completed one of two lists with one version of each item truncated before the critical word. They were instructed to complete the sentence with the first sensible word coming to mind. Cloze value was computed as the percentage of participants who used the intended critical word.

We subsequently determined whether sentences (truncated after the critical word) were, on average, regarded as true or false. Twenty-four different students evaluated one of four counterbalanced sentence lists containing only one condition per quadruplet, and decided whether the sentences were true (1 = False, 7 = True), skipping any they could not evaluate.

Based on these results, we excluded quadruplets with low cloze value, containing true/false sentences rated below/over 4, or sentences skipped by more than two participants. In the ultimate set of 120 quadruplets, true and false sentences had similar cloze values and ratings across conditions (see Table 1), and critical words were matched for mean log

frequency (CT/RWT = 1.44/1.55; $p = .11$; Davis & Perea, 2005) and word length (CT/RWT = 6.65/6.89 letters; $p = .24$). Average counterfactual sentence length was 16.2 words ($SD = 2.8$), and real-world sentence length was 14.6 words ($SD = 2.7$).

We created four counterbalanced lists so that each sentence appeared in only one condition per list, but in all conditions equally often across lists. Within each list, items were pseudo-randomly mixed with 60 filler sentences to limit succession of identical sentence types while matching trial types on average list position. The filler sentences did not start with ‘Si’ or ‘Como’, and consisted of two clauses separated by a comma.

Experimental procedure

While inside the scanner, participants read sentences presented via back-projection onto the middle of the screen, and would view the stimuli via a mirror attached to the head coil. They were instructed to minimize movement and read the sentences attentively, and to answer yes/no comprehension questions that appeared after some of the sentences with a left- or right-hand button-press (left- or right-hand assignment for yes/no was counterbalanced across participants).

Four trial lists were used (each subject was pseudorandomly assigned to one of the four trial lists, so that the lists were equally distributed across subjects). For the first list, 30 items from each condition were pseudo-randomly mixed with the filler sentences such that no trial type occurred more than three times consecutively and trials of each type were matched on average list position. The other lists were derived from the first by rotating the trial types. The total of 180 sentences was divided into 6 sessions (presented in fixed-order across trial lists) of approximately 8 minutes each. Following the third session, participants exited the scanner for a short break.

Participants silently read sentences presented in black letters on a light-grey background. The first clause was presented as a whole for 4000 ms, followed by a fixation cross and blank screen each for 500 ms, and the second clause was presented word-by-word (screen-centered, 400 ms word duration, 200 ms inter-word-interval). Following every final word, a blank screen was presented for 500 ms, followed by either a fixation period of variable duration (5-8 seconds) or by first a yes/no comprehension question and then a fixation period. During the fixation period, participants fixate on a cross in the middle of the screen and awaited the start of the next trial. If a comprehension question appeared, it was presented for 2 seconds, followed by a response screen that disappeared upon button-press or after 1 second. The questions only served to keep participants more attentive. They always probed their world knowledge related to words in the sentence (e.g., “Does N.A.S.A. have a headquarters in Houston?”) but did not probe knowledge that was critical for understanding the sentences, and did not probe about the counterfactual premise. These 60 questions (30 requiring a ‘yes’ button-press response) were distributed across sentence types. Participants performed with 92 % accuracy on average (range across subjects 77-98%).

fMRI data acquisition, preprocessing and analysis

Imaging took place on a 3-T MR scanner (Siemens TrioTim) with echoplanar imaging capability. Head motion was minimized using pillows and cushions around the head. Each subject then viewed one of the four counterbalanced sentence lists with the sentence trials and fixation trials, across six functional runs. Each functional run lasted around 570 s during which whole head T2^{*}-weighted EPI-BOLD fMRI data were acquired using an interleaved even acquisition EPI sequence (TR = 2 s; TE = 30ms; flip angle = 90°; 32 axial slices; matrix size = 64×64; slice thickness = 3 mm; slice gap = 0.75 mm;

transverse orientation acquisition; isotropic voxel-size = $3 \times 3 \times 3 \text{ mm}^3$). After the functional runs, subjects underwent one high-resolution 3D structural scan, using a T1-weighted MPAGE sequence (176 transverse slices; volume TR = 2530 ms; TE = 2.97 ms; TI = 1100 ms; transverse orientation acquisition; flip angle = 7° ; slice matrix = 256×256 ; slice thickness = 1 mm, slice gap = 0.5 mm).

Image preprocessing and statistical analysis was performed using the SPM8 software (<http://www.fil.ion.ucl.ac.uk>). The functional EPI-BOLD contrast images were realigned, and the mean of realigned images was co-registered with the corresponding structural MRI by using mutual information optimization. These images were subsequently slice-time corrected, spatially normalized (images were re-sampled with a $2 \times 2 \times 2 \text{ mm}^3$ resolution), transformed into a common space (MNI-T1 template), and spatially filtered with an isotropic 3D Gaussian kernel (10 mm FWHM). The data were analyzed using the general linear model and statistical parametric mapping. We included the following explanatory variables (see Table 2, for a schematic representation): each of the four critical conditions (1 = Counterfactual True, 2 = Counterfactual False, 3 = Real-world True, 4 = Real-world True) modeled separately from the onset of the critical word with a duration extending to the end of the sentence including the sentence-final word, counterfactual context modeled from the onset of each counterfactual sentence with a duration extending to the offset of the pre-critical word (pooled across counterfactual true and false sentences because they were identical up to the critical word, number 5 in Table 2), real-world context modeled from the onset of each real-world sentence with a duration up extending to the offset of the pre-critical word (pooled across real-world true and false sentences because they were identical up to the critical word, number 6 in Table 2). Importantly, pre-

critical word differences between counterfactual and real-world sentences in presence/absence of negation and in sentence length can thus not account for observed differences in the critical regions. The fixation period was modeled from the onset of the fixation mark until it disappeared from the screen. Effects of no-interest included one regressor that pooled all filler sentences (modeled from sentence onset to offset), one regressor for the comprehension question and response screen, and additional regressors for session and subject effects. The explanatory variables were modeled as a fixed response (box-car) waveform temporally convolved with the canonical HRF along with its temporal derivative (Friston et al., 1998), while controlling for serial correlations. Low-frequency noise was removed with a high-pass filter (time constant 128 s).

Table 2. Schematic representation of the regressors corresponding to the different sentence parts		
Counterfactual Context	5. If N.A.S.A. had not developed its Apollo Project, the first country to land on the moon would have been	1. Russia surely.
	5. If N.A.S.A. had not developed its Apollo Project, the first country to land on the moon would have been	2. America surely.
Real-World Context	6. Because N.A.S.A. developed its Apollo Project, the first country to land on the moon was	3. America surely.
	6. Because N.A.S.A. developed its Apollo Project, the first country to land on the moon was	4. Russia surely.

Region-of-interest analysis

Given the a priori hypothesis about the role of the IFG, a region-of-interest (ROI) analysis was performed using the Marsbar toolbox (Brett et al., 2002) by extracting average

parameter estimates per condition and per subject for 2 LIFG ROIs and their right hemisphere counterparts. These ROIs were based on the results of a related study that reported activations within different subregions of the left and right IFG (pars orbitalis/triangularis; BA 45/47) for world knowledge violations in healthy adults during sentence processing (Tesink et al., 2011; see also Tesink et al., 2008, for similar results)¹. Signal was sampled from 2 spherical ROIs with a 10 mm radius centered on coordinates [-44 29 12] and [36 26 -8] for BA 45 and 47 respectively (see Figure 1a). These ROI center coordinates approximated the average of peak voxel coordinates reported by Tesink et al (2011) (BA 45: [-40 30 4] and [-44 28 14], BA 47: [38 22 -10] and [-34 30 -8]), with a small adjustment to avoid ROI overlap. A 2(Context: counterfactual, real-world) by 2(Propositional truth-value: true, false) by 2(hemisphere: left, right) repeated measures

¹ The coordinates were based on the results reported by Tesink et al. (2011) rather than those reported by Hagoort et al. (2004) or Menenti et al. (2009). Hagoort et al. reported coordinates that referred to the common activation for semantic violations and world knowledge violations compared to the correct condition, and may thus not be optimal for investigating the current effects of sentence truth-value (for example, semantic violations and world knowledge violations are associated with different peak activity coordinates, as reported by Tesink et al.). In addition, while Hagoort et al. only reported a single cluster of common activation in the LIFG, the Tesink et al. study revealed bilateral effects of truth-value for the same items. Moreover, while Tesink et al. reported effects of truth-value of single sentences, the Menenti et al. study involved multi-sentence stories rather than single sentences, and the reported coordinates in that study might thus not be optimal for the comparisons in the current study.

analysis of variance (ANOVA) was performed on the extracted parameter estimates per ROI.

Whole-brain analysis

In addition to ROI analyses, a whole-brain analysis was performed to examine whether other brain regions were sensitive to counterfactual and real-world truth-value. Average parameter estimates for the explanatory variables were generated for each subject, and subjected to a second-level random effects analysis with non-sphericity correction for correlated repeated measures, according to the 2 (Context: counterfactual, real-world) by 2 (Truth-value: true, false) design. The results of the random effects analyses were thresholded at $P = 0.001$ (uncorrected) and the cluster-size statistics were used as the test statistic, only clusters are reported that were significant at $P \leq 0.05$ corrected for multiple comparisons using the false discovery rate (FDR; Genovese, Lazar & Nichols, 2002). All local maxima are reported as MNI coordinates (Evans et al., 1993). Anatomical location and approximate Brodmann areas were determined using the AAL toolbox for SPM8 (Tzourio-Mazoyer et al., 2002) and with the xjView toolbox (www.alivelearn.net/xjview8).

RESULTS

ROI analysis

Differences in parameter estimates corresponding to the effect of truth-value (false minus true sentences) per sentence type and ROI are plotted in Figure 1b. Overall, false sentences elicited more activity than true sentences in the BA 45 ROIs ($F(1,23) = 31.43, p < .001$) and in the B47 ROIs ($F(1,23) = 80.98, p < .001$). In the BA 45 ROIs the effect of truth-value was similar for counterfactual and real-world sentences ($F(1,23) = .005, p =$

.95), and did not differ between the left and right ROI ($F(1,23)= 1.11, p = .30$). In contrast, in the BA 47 ROIs the effect of truth-value was larger for counterfactual sentences than for real-world sentences ($F(1,23)= 5.08, p = .034$), but this pattern differed for the left and right ROI, as reflected in a significant 3-way interaction effect ($F(1,23)= 4.58, p = .043$). The effect of truth-value was greater for counterfactual sentences than for real-world sentences in the right BA47 ROI ($F(1,23)= 6.90, p = .015$), but not in the left BA 47 ROI ($F(1,23)= .13, p = .73$). This differential effect of truth-value in the right BA 47 ROI for the two sentence types was of a quantitative nature rather than a qualitative: the effect of truth-value was statistically significant for each sentence type, but it was stronger in counterfactual sentences ($F(1,23)= 6.83, p < .001$) than in real-world sentences ($F(1,23)= 2.21, p = .037$). This difference was driven mainly by a large response to counterfactual false sentences, which was larger than the response to real-world false sentences ($F(1,23)= 3.59, p = .002$), while the responses to counterfactual true and real-world true sentences did not differ ($F(1,23)= 1.06, p = .31$).

Whole-brain analysis

The whole-brain analysis did not reveal any significant clusters for the interaction between Context and Propositional truth-value. However, because the ROI analyses revealed differential effects of truth-value in counterfactual and real-world sentences, the following linear contrasts (and their reverse counterparts) were further specified: Counterfactual False > Counterfactual True, Real-world False > Real-world True, Counterfactual > Real-world. Consistent with the absence of a significant Context by Propositional truth-value interaction effect, the false minus true contrast elicited similar activation increases for counterfactual and real-world sentences (see Table 3 and Figure 2).

These clusters were located in left and right inferior frontal gyrus (LIFG and RIFG, BA 45 and BA 47) extending into the middle frontal gyrus (MFG, BA 8/9), left middle temporal gyrus (MTG, BA 21), medial parts of the superior frontal gyrus (SFG, BA 6/8), and the left inferior parietal lobule (IPL, BA 40). For real-world sentences an additional cluster was located in the caudate nucleus. These results are a clear replication of effects of world-knowledge violations as reported by Hagoort and colleagues (Groen et al., 2010; Hagoort et al., 2004; Menenti et al., 2009; Tesink et al., 2009, 2011).

Finally, the reverse contrast yielded no clusters that showed more activity to true sentences than to false sentences. However, there was a main effect of Context: counterfactual sentences overall elicited more activity than real-world sentences in the left and right middle temporal gyrus, a finding that is consistent with the results of Menenti et al. (2009), who reported more activation in this region for counterfactual context compared to real-world context. Consistent with the main effect of context, the contrast (Counterfactual False > Real-world False) elicited a significant cluster in LMTG ($k = 440$, peak voxel coordinates $[-48 -32 4]$), as did the contrast (Counterfactual True > Real-world True) ($k = 1421$, peak voxel coordinates $[-50 -24 -14]$). Consistent with the interaction effect observed in the ROI analysis, the contrast (Counterfactual False > Real-world False) elicited a significant cluster in RIFG ($k = 837$, peak voxel coordinates $[36 28 -2]$), whereas the contrast (Counterfactual True > Real-world True) did not.

Table 3.							
						MNI coordinates	
Anatomical region	BA	<i>p</i> (Cluster)	Cluster size	Voxel <i>T</i> Value	<i>x</i>	<i>y</i>	<i>z</i>
<i>a. Results from the Whole-brain analysis for the contrast Counterfactual False > Counterfactual True*</i>							
LIFG (pars orbitalis)	47	.000	4765	6.92	-36	22	-6
	47			6.62	-44	26	-10
LIFG (pars triangularis)	45			6.57	-40	22	4
RIFG (pars orbitalis)	47	.000	4198	6.86	32	26	-4
R Middle frontal gyrus	9			6.33	52	20	40
RIFG (pars orbitalis)	47			6.26	46	30	-10
Medial Superior frontal gyrus	6	.000	3658	5.97	-6	22	56
	8			5.96	2	38	50
L Superior frontal gyrus	6			5.95	-14	12	58
L Middle temporal gyrus	21	.000	1083	5.49	-48	-34	-4
				4.98	-52	-26	-8

					4.17	-64	-48	4
L Inferior parietal lobule	40	.000	968		4.98	-54	-52	48
					4.76	-50	-48	38
					3.88	-64	-46	40
<i>b. Results from the Whole-brain analysis for the contrast Real-world False > Real-world True</i>								
LIFG (pars orbitalis)	47/11	.000	5631		7.21	-46	46	-14
L Middle frontal gyrus	8				6.04	-50	14	44
LIFG (pars triangularis)	45				5.75	-50	24	-2
L Middle temporal gyrus	21	.000	886		6.72	-50	-24	-12
L Inferior parietal lobule	40	.000	2085		5.90	-48	-50	34
					5.81	-54	-52	48
					4.60	-38	-56	46
Medial Superior frontal gyrus	9	.000	2525		5.11	-10	36	32
L SMA	8				4.94	-2	22	58
L Superior frontal gyrus	6				4.83	-14	26	38
RIFG (pars orbitalis)	47/11	.000	2444		4.99	46	46	-12

R Middle frontal gyrus	8				4.91	42	18	38
	9				4.57	36	16	30
R Inferior parietal lobule	40	.003	557		4.60	48	-56	42
R Angular gyrus	40				4.09	46	-52	34
R Caudate		.006	456		4.13	10	4	14
L Caudate					3.89	-10	8	10
<i>c. Main effect of context (Counterfactual > Real-world)</i>								
L Middle temporal gyrus	21	.000	2365		6.10	-50	-24	-14
					5.71	-48	-32	4
					5.01	-52	-2	-24
R Middle temporal gyrus	21	.064	338		4.73	52	-18	-18
					4.38	44	-16	-14
					3.94	50	-8	-12
* <i>P</i> -values correspond to cluster-level statistical tests with FDR correction at $P \leq 0.05$. <i>T</i> -values correspond to the local maxima (multiple local maxima are reported when they are more than 8 mm apart).								

DISCUSSION

This fMRI study investigated the neural circuits that are sensitive to the propositional truth-value of sentences about counterfactual worlds or about the real world, and aimed to reveal hemispheric differences between the inferior prefrontal gyri in processing counterfactual truth-value and real-world truth-value. BOLD responses were compared to critical words that rendered counterfactual or real-world sentences true or false, i.e. in a 2 (Context: counterfactual, real-world) by 2 (Truth-value: true, false) design. The main result of this study, as evident from ROI analyses involving the left and right BA 45 and 47, was that whereas the left BA 47 showed similar activity increases to counterfactual false and to real-world false sentences (compared to true sentences), the RIFG BA 47 showed a larger increase for counterfactual sentences than for real-world sentences². No such pattern was found for the left and right BA 45 ROIs. The second contribution of this study is that, across both sentence types, false sentences elicited

² Whereas the counterfactual sentences and real-world sentences also differed in sentence length and the presence of negation, these differences do not explain the currently observed effects. First of all, these differences occurred before the critical words, and their effects were accounted for by separate. The part of the sentence that preceded the critical words was modeled separately from the part of the sentence from the critical word onwards. Because this part of the sentence was identical for the true and false sentences, this regressor pooled the data from the true and false sentences (see Table 2). Importantly, activity associated with these differences is thus accounted for in the model in a way that it cannot explain the observed effects of the critical words. The regressors that correspond to the critical sentence parts involve exactly the same words for the four conditions. Secondly, the critical comparisons in the ROI analysis showed that there were no differences between the counterfactual true and real-world true conditions, suggesting that these differences did not affect the observed results.

activation increases compared to true sentences in several brain regions linked to semantic processing (left and right IFG, superior and middle frontal gyrus, middle temporal gyrus, and inferior parietal cortex; Bookheimer, 2002; Hagoort et al., 2009; Lau et al., 2008; Price, 2010).

The fact that the RIFG was more sensitive to counterfactual truth-value than to real-world truth-value is consistent with greater RIFG sensitivity to counterfactual discourse coherence (Menenti et al., 2009). However, it is important to note that the hemispheric differences for processing counterfactual truth-value and real-world truth-value emerged as a function of the LIFG and RIFG BA 47 responses to real-world truth-value, whereas these regions showed similar effects of counterfactual truth-value. Larger effects of truth-value of ‘simple’ sentences in LIFG compared to RIFG have already been reported in similar studies, along with stronger effects of lexical-semantic violations in LIFG than in RIFG (e.g., Groen et al., 2010; Hagoort et al., 2004; Tesink et al., 2009). The current larger sensitivity of the RIFG to counterfactual truth-value than to real-world truth-value may thus reflect increased semantic processing due to the semantic complexity of the counterfactual false sentences. Whereas real-world false sentences increase semantic processing associated with rejecting the sentence as being false based on factual knowledge, counterfactual false sentences may elicit extended semantic processing to reject the sentence as being false (e.g., America would probably not be the first to land on the moon if it hadn’t been for N.A.S.A.) despite the fact that the sentence is congruent with actual facts about the world (e.g., America did land on the moon first). The larger effect in the RIFG for counterfactual truth-value is consistent with right hemisphere activations associated with comprehending contextual and figurative meaning and with complex sentences and discourse level

processing (e.g., Kuperberg, Lakshmanan, Caplan & Holcomb, 2006; Stringaris et al., 2006; Xu et al., 2006 see Beeman, 2005; Mason & Just, 2006). On one interpretation, this effect may reflect intensified semantic selection (i.e., inhibiting competing concepts to select one concept for message-level semantic integration) to process information that is activated in the right hemisphere (e.g., Jung-Beeman, 2005). Perhaps that real-world consistent but contextually irrelevant concepts ('America') are more strongly activated in the right hemisphere, and that falsifying counterfactual sentences in face of these 'lures' therefore engages more processing resources in this hemisphere.

Whereas in the current study counterfactual false and real-world false sentences similarly elicited widespread effects, Menenti et al. (2009) reported a reduced effect of world knowledge in the IFG following the counterfactual context, and while this reduction was greater in the RIFG than in the LIFG, the effect of world knowledge in the RIFG was not reversed by counterfactual context. These differences in the effects of counterfactual context may be explained by the difference in counterfactual constraints in the two studies. Counterfactual true continuations in the current study were, on average, as expected as real-world true continuations, and they received similar truth-value ratings. In contrast, the counterfactual discourse context used by Menenti and colleagues made the counterfactual continuation less unexpected (43% compared to 0% following the real-world context) yet the real-world continuation remained to be relatively expected too despite the counterfactual context (34% compared to 77% following the real-world context). Thus, absence of a counterfactual context effect in the RIFG may reflect the fact that world knowledge violations and correct sentences were equally unexpected and/or plausible. In the current study, one factor that may have contributed to making counterfactual contexts as

‘strong’ as the real-world contexts may have been the use of historical counterfactual conditionals. Whereas Menenti et al created novel and unfamiliar fictional stories, alternative endings to known historical events may be more easily computed, for example because relevant information is also part of our real-world knowledge (e.g., the fact that the Soviets were also making substantial progress in landing somebody on the moon at the time that the USA managed to do so). Theories of counterfactual comprehension indeed assume that the more dissimilar a counterfactual world is to the real world the more difficulty people have in reasoning about them (e.g., Lewis, 1973; see Byrne, 2007, for discussion).

The current findings are also in line with recent ERP findings that suggest that, rather than being ‘message-blind’, the right hemisphere is sensitive to discourse-level comprehension (e.g., Coulson, Federmeier, Van Petten & Kutas, 2005; Federmeier, Wlotko & Meyer, 2008). In an earlier ERP version of the current study, counterfactual false sentences and real-world false sentences elicited identical N400 effects compared to their respective true version, reflecting the fact that incoming words are being mapped onto the counterfactual context without any delay (Nieuwland & Martin, 2011). Although different participants and dissimilar techniques limit generalization across the two sets of results, these results combined may suggest that the increased RIFG response does not reflect the early semantic processes that are thought to be indexed by the N400 ERP. This is consistent with recent claims in the literature that neural sources underlying N400 modulations lie in temporal lobes rather than inferior prefrontal regions (see Van Petten & Luka, 2006; Lau et al., 2008). In the current study, counterfactual and real-world sentences elicited similar effects of truth-value in the left middle temporal gyrus. Importantly, fMRI has a low temporal resolution and in the current study the condition regressors were modeled from

critical word onset to sentence offset, picking up on effects outside the critical-word time-locked N400 time window. Therefore, the current RIFG effects reflect differences between processing of counterfactual truth-value and real-world truth-value that occur at a wider timescale.

The extended pattern of activations that was associated with false sentences independently replicates BOLD responses to world knowledge violations as reported by Hagoort and colleagues (Groen et al., 2010; Hagoort et al., 2004; Menenti et al., 2009; Tesink et al., 2009, 2011). Moreover, the current study generated a similar pattern of results for counterfactual and real-world sentences, despite the fact that effects of propositional truth-value for these sentence types were computed over opposite pairs of critical words (thereby effectively controlling for lower-level lexical-associate factors that are also known to modulate activity in inferior frontal gyri, see Lau et al., 2008). Consistent with earlier studies, false sentences induced increased demands on a wider semantic processing network even when participants are not explicitly instructed to determine truth-value. These results may reflect increased semantic retrieval of relevant information in long-term memory (e.g., of the correct information) from medial temporal lobe, governed by the inferior prefrontal gyri (e.g., Bookheimer, 2002; Hagoort et al., 2009; Lau et al., 2008). This is also consistent with the extensive structural and functional connectivity as observed for within the fronto-temporal language network and for connectivity for regions in this network with dorsolateral and inferior parietal regions (e.g., Turken & Dronkers, 2011; Xiang, Fonteijn, Norris & Hagoort, 2010). The specific role of these other regions in the reported network remains elusive. The inferior parietal lobule may support more general reprocessing demands, e.g. increased demands on verbal working memory (e.g., Jonides et al., 1998).

Interestingly, activity increases in this region have been reported for syntactic errors but not for semantic anomalies (they may even result in relative deactivations; e.g., Hagoort et al., 2004; Nieuwland, Petersson & Van Berkum, 2007; Nieuwland, Martin & Carreiras, 2011; Tesink et al., 2009). Sentences that are false or grammatically incorrect have in common that the reader or listener may automatically tap into what the correct (i.e., true or grammatical) sentence could be, whereas there may not be such a straightforward solution to deal with semantically odd sentences. In contrast to the IPL, the medial superior frontal and middle frontal gyrus also sometimes show activity increases to semantic violations. In the MUC model of language (e.g., Hagoort et al., 2009), these approximate regions map onto the Control component that deals with the fact that the language system must operate in a communicative context (e.g., Willems et al., 2010). Future research can elucidate whether these neural circuits that are sensitive to propositional truth-value are modulated by the ease with which truth-value can be established, or by the relevance of propositional truth-value to the unfolding discourse context.

In sum, counterfactual false sentences elicit greater activation increases in RIFG than matched real-world false sentences, but not in LIFG, consistent with greater involvement of the RIFG in counterfactual discourse comprehension. These results reflect differential hemispheric sensitivity to propositional truth-value of sentences about counterfactual worlds or about the real world. Moreover, whole-brain analyses showed that counterfactual false and real-world false sentences both elicited activation increases in left and right inferior prefrontal gyrus, middle and superior frontal gyrus, left middle temporal gyrus and in the inferior parietal lobule, revealing a distributed neural circuit for dealing with propositional truth-value.

LITERATURE

- Badre, D., & Wagner, A.D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews*, 1(3), 206-218.
- Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, 25, 151–188.
- Braine, M. D. S., & O'Brien, D. P. (1991). A theory of *If*: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, 98, 182–203.
- Brett, M, Anton, J., Valabregue, R., & Poline, J. (2002). Region of interest analysis using the MarsBar toolbox for SPM 99. Poster presented at the 8th International Conference on Functional Mapping of the Human Brain. Sendai, Japan.
- Byrne, R.M.J. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Sciences*, 6(10), 426-431.
- Byrne, R.M.J. (2007). Précis of the rational imagination: How people create alternatives to reality. *Behavioral and Brain Sciences*, 30, 439–480.
- Byrne, R.M.J. & Johnson-Laird, P.N. (2009). 'If' and the problems of conditional reasoning. *Trends in Cognitive Sciences*, 13(7), 282-287.
- Coulson, S., Federmeier, K.D., Van Petten, C., Kutas, M. (2005). Right hemisphere sensitivity to word- and sentence-level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 129–147.
- Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews Neuroscience*, 1, 41-50.
- Evans, J. S. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review*, 13(3), 378-395.
- Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L. & Peters, T. M. (1993). *IEEE Conference Record, Nuclear Science Symposium, and Medical Imaging Conference* (Inst. Electr. Electron. Eng., San Francisco), pp. 1813–1817.
- Evans, J.S.B.T., Over, D.E. & Handley, S.J. (2005). Suppositions, extensionality and conditionals: a critique of the model theory of Johnson-Laird & Byrne (2002). *Psychological Review*, 112, 1040-1052.

- Faust, M.E., Gernsbacher, M.A. (1996). Cerebral mechanisms for suppression of inappropriate information during sentence comprehension. *Brain and Language*, 53, 234–259.
- Federmeier, K. D., Wlotko, E., Meyer, A. M. (2008). Whats right in language comprehension: ERPs reveal right hemisphere language capabilities. *Language and Linguistics Compass*, 2, 1-17.
- Ferguson, H.J., & Sanford, A.J. (2008). Anomalies in real and counterfactual worlds: An eye-movement investigation. *Journal of Memory and Language*, 58, 609-626.
- Ferguson, H.J., Sanford, A.J., & Leuthold, H. (2008). Eye-movements and ERPs reveal the time-course of processing negation and remitting counterfactual worlds. *Brain Research*, 1236, 113-125.
- Ferstl, E.C., Neumann, J., Bogler, C., and von Cramon, D.Y. (2008). The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, 29, 581-593.
- Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *Neuroimage*, 7(1), 30-40.
- Genovese, C.R., Lazar, N.A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15(4), 870-878.
- Gernsbacher, M. A. (1997). Attenuating interference during comprehension: The role of suppression. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp. 85-104). San Diego, CA: Academic Press.
- Goel V. (2007). Anatomy of deductive reasoning. *Trends in Cognitive Sciences*, 11(10):435–441.
- Goel, V., & Dolan, R.J. (2003). Explaining modulation of reasoning by belief. *Cognition*, 87(1), B11-B22.
- Groen, W. B., Tesink, C. M. J. Y., Petersson, K. M., Van Berkum, J. J. A., Van der Gaag, R. J., Hagoort, P., & Buitelaar, J. K. (2010). Semantic, factual, and social language comprehension in adolescents with autism: An fMRI study. *Cerebral Cortex*, 20(8), 1937-1945.
- Jonides, J., Schumacher, E.H., Smith, E.E., Koeppe, R.A., Awh, E., Reuter-Lorenz, P.A.,

- Marchuetz, C. & Willis, C.R. (1998) The role of parietal cortex in verbal working memory. *Journal of Neuroscience*, 18, 5026–5034.
- Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, 9(9), 416–423.
- Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences*, 4th ed. (pp. 819-836). Cambridge, MA: MIT Press.
- Hagoort, P., Hald, L.A., Bastiaansen, M. & Petersson, K.M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304, 438-441.
- Hagoort, P., & van Berkum, J. J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society, Series B*, 362, 801–811.
- Hald, L.A., Steenbeek-Planting, E.G., & Hagoort, P. (2007). The interaction of discourse context and world knowledge in online sentence comprehension. Evidence from the N400. *Brain Research*, 1146, 210-218.
- Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, 9(11), 712-718.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182.
- Kuperberg, G.R., Lakshmanan, B.M., Caplan, D.N., & Holcomb, P.J. (2006). Making sense of discourse: an fMRI study of causal inference across sentences. *NeuroImage*, 33, 343-361.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De) constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933.
- Lewis, D. (1973). *Counterfactuals*. Cambridge, Mass.: Harvard University Press.
- Mason, R.A., and Just, M.A. (2006). Neuroimaging contributions to the understanding of discourse processes. In M. Traxler and M.A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (pp. 765-799). Amsterdam: Elsevier.
- Menenti, L., Petersson, K. M., Scheeringa, R., & Hagoort, P. (2009). When elephants fly: Differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. *Journal of Cognitive Neuroscience*, 21, 2358-2368.
- Montague, R., (1974). The Proper Treatment of Quantification in Ordinary English. In

- Thomason R. (ed.), *Formal Philosophy: Selected Papers of Richard Montague*, Yale University Press, New Haven.
- Myers, J.L. & O'Brien, E.J., (1998). Accessing the discourse representation during reading. *Discourse Processes*, 26, 131-157.
- Nieuwland, M.S. & Kuperberg, G.R. (2008). When the truth is not too hard to handle: An event-related potential study on the pragmatics of negation. *Psychological Science*, 19, 1213-1218.
- Nieuwland, M. S., & Martin, A. E. (2011). If the real world were irrelevant, so to speak: The role of propositional truth-value in counterfactual sentence comprehension. *Cognition*, doi:10.1016/j.cognition.2011.09.001.
- Nieuwland, M.S., Martin, A.E., & Carreiras, M. (2011). Brain regions that process case: Evidence from Basque. *In press at Human Brain Mapping*.
- Nieuwland, M.S., Petersson, K.M. & Van Berkum, J.J.A. (2007). On sense and reference: Examining the functional neuroanatomy of referential processing. *NeuroImage*, 37(3), 993-1004 .
- Nieuwland, M.S. & Van Berkum, J.J.A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, 18(7), 1098-1111.
- Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences* 13 (4): 707-784.
- Price, C.J., (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, 1191 (1), 62–88.
- Searle, J.R. (1975). The logical status of fictional discourse. *New Literary History*, 6, 319-332.
- Stalnaker, R.C. (1968). A theory of conditionals. In Rescher, N. (ed.), pages 98–112. *American Philosophical Quarterly Monograph Series*. Blackwell, Oxford.
- Stringaris, A. K., Medford, N., Giora, R., Giampietro, V. C., Brammer, M. J., & David, A. S. (2006). How metaphors influence semantic relatedness judgments: The role of the right frontal cortex. *Neuroimage*, 33, 784–793.
- Tarski, A. (1944). The Semantic Conception of Truth. In Martinich, A.P. (ed.), pages 69-91. *The Philosophy of Language*. New York: Oxford University Press.

- Tesink, C. M. J. Y., Buitelaar, J. K., Petersson, K. M., Van der Gaag, R. J., Teunisse, J.-P., & Hagoort, P. (2011). Neural correlates of language comprehension in autism spectrum disorders: When language conflicts with world knowledge. *Neuropsychologia*, 49, 1095-1104.
- Tesink, C. M. J. Y., Petersson, K. M., Van Berkum, J. J. A., Van den Brink, D., Buitelaar, J. K., & Hagoort, P. (2009). Unification of speaker and meaning in language comprehension: An fMRI study. *Journal of Cognitive Neuroscience*, 21, 2085-2099.
- Thompson-Schill S.L., D'Esposito M., Aguirre G.K., Farah M.J., (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences, USA*, 94:14792–14797.
- Turken, A.U., & Dronkers, N.F. (2011). The Neural Architecture of the Language Comprehension Network: Converging Evidence from Lesion and Connectivity Analyses. *Frontiers in Systems Neuroscience*, 5:1. doi: 10.3389/fnsys.2011.00001.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15, 273–289.
- Van Berkum, J. J. A., Holleman, B., Nieuwland, M. S., Otten, M., & Murre, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science*, 20, 1092 -1099.
- Van Berkum, J. J. A., Van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, 20(4), 580-591.
- Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, 97(3), 279–293.
- Willems, R. M., De Boer, M., De Ruiter, J.P., Noordzij, M. L, Hagoort, P. & Toni, I. (2010). A cerebral dissociation between linguistic and communicative abilities in humans. *Psychological Science*, 21(1), 8-14.
- Xiang, H.-D., Fonteijn, H. M., Norris, D. G., & Hagoort, P. (2010). Topographical functional connectivity pattern in the perisylvian language networks. *Cerebral Cortex*, 20, 549-560.

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *Neuroimage*, 25,1002–1015.

ACKNOWLEDGEMENTS

This research was supported by a Plan Nacional research grant from the Spanish Ministry of Science and Innovation (grant number PSI2010-18087) to MSN. Many thanks to Javi Miqueleiz for help with stimulus construction, to Natalia Barrios and Larraitz Lopez for help with data collection, and to two anonymous reviewers and Andrea Eyleen Martin for comments on an earlier version of this manuscript.

FIGURE CAPTIONS

Figure 1. (a) ROIs in the current study. Two 10-mm sphere centered at MNI coordinates [-44 29 12] (BA 45, top) and [-36 26 -8] (BA 47, bottom) and their right hemisphere equivalents. (b) Effect of propositional truth-value in counterfactual sentences (dark grey bars, ‘CF’) and real-world sentences (light gray bars, ‘RW’) in four ROIs (left/right, BA 45/47). Displayed are the false-minus-true difference score (and 95 % confidence intervals) in average beta parameter value per sentence type and per ROI. (* $p < .05$)

Figure 2. Effect of propositional truth-value in counterfactual sentences (a) and in real-world sentences (b). Pair-wise comparisons for false minus true sentences across all subjects (thresholded at $P \leq 0.001$ uncorrected). Labels are provided only for significant clusters (corrected for multiple comparisons with a cluster-level FDR $P \leq 0.05$). LIFG = Left inferior frontal gyrus. RIFG = Right inferior frontal gyrus. SFG = Superior frontal gyrus. MFG = Middle frontal gyrus. MTG = Middle temporal gyrus. IPL = Inferior parietal lobule.

Figure 1
[Click here to download 9. Figure: Figure_1.pdf](#)

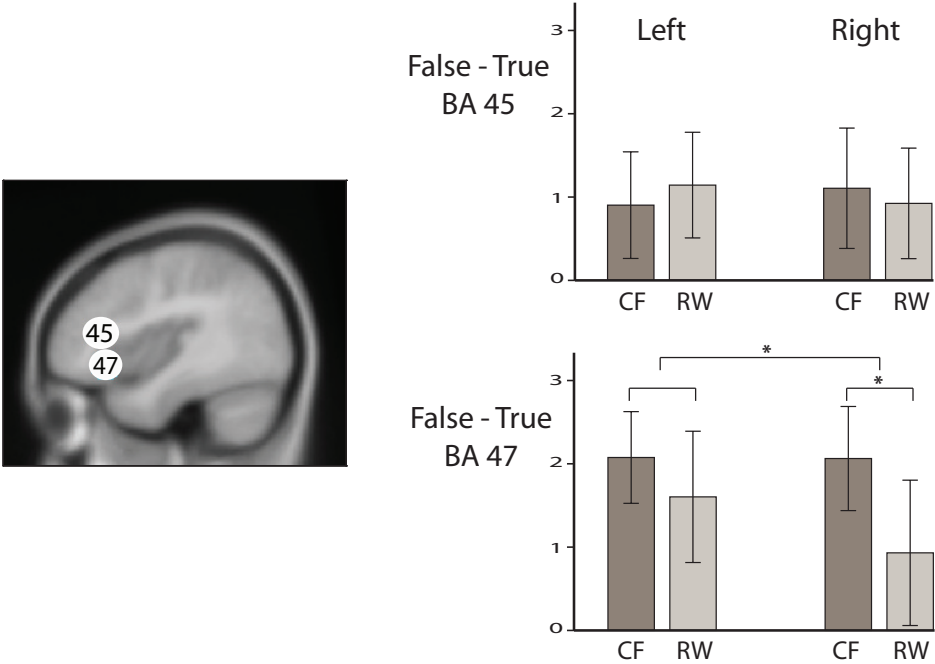


Figure 2
[Click here to download 9. Figure: Figure_2.pdf](#)

